

Translation of H01-209500

Fig. 3 illustrates how the synthesis lattice is formed in the embodiment of the present invention.

Fig. 3 illustrates a part of the synthesis unit lattice, which is formed when a mark of a series of the phonological sound, /hanagasakidashita/, is entered using the synthesis unit entry file shown in Fig. 2. In this example, the longest unit entry starting with /h/, /hana/, is first found out through the search of the entry file. The unit entries shorter than /hana/ including /h/, /ha/, and /han/ are also found out in this search process. The information stored in the entry file as shown in Fig. 3 can be thus read out. Therefore, it is found that /hana/ has 10 templates including /hana/ ("flower"), /hana/ ("nose") and /hanayaka/ ("gorgeous") as the word speech samples.

Fig. 4 illustrates speech units selected in the synthesis unit lattice of s based on certain assumed selection criteria in the embodiment of the present invention as well as word speeches including templates of the speech units.

In Fig. 4, four criteria are assumed to determine the speech template for the entered example as shown in Fig. 3, which include (1) priority to continuity of transition parts from a consonant to a vowel, (2) priority to continuity of voiced sound parts, (3) priority to a long unit, and (4) priority to connectionability between units. As shown in the drawing, /hana/ is obtained from the beginning of /hanabanashii/ ("gorgeous") as the word speech sample. Also, /ga/, /sa/, /kida/, and /shita/ are obtained from /nagasa/ ("length"), /murasaki/ ("purple"), /hikidashi/ ("drawer") and /ashita/ ("tomorrow") respectively. As a specific method to achieve selection based on these criteria, the selection method in the conventional morpheme identification method for language processing is applicable through costing the selection according to the priority of each of the selection criteria as shown in the article titled "Word Identification

Method for Japanese Text-to-Speech Conversion System" (Proc.
of ICASSP 86, Page 2411-2414), disclosed in April 1986 by
Yoshinori Sagisaka and Hirokazu Sato.

SPEECH SYNTHESIS SYSTEM

1

Patent Number: JP1209500
Publication date: 1989-08-23
Inventor(s): KOSAKA YOSHINORI
Applicant(s): A T R JIDO HONYAKU DENWA KENKYUSHO:KK
Requested Patent: JP1209500
Application Number: JP19880036204 19880217
Priority Number(s):
IPC Classification: G10L5/04
EC Classification:
Equivalents: JP2880507B2

Abstract

PURPOSE:To select a proper speech unit rationally and efficiently by storing previously a group of all speech unit candidates which are used possibly with respect to a phoneme sequence symbol signal inputted in a speech unit set.

CONSTITUTION:A synthetic unit entry file 4 stores various pieces of speech synthesis unit information which are required for speech synthesis. A synthesis unit lattice generation part 3 finds phonemes which are stored previously and present as synthesis units in a partial phoneme sequence outputted from an input phoneme sequence processing part 2 by retrieving the synthesis unit entry file 4, and lists and supplies all speech synthesis unit candidates which are used possibly for synthesis to a speech template determination part 5. Then, speech units which are used according to a prescribed standard are determined in the speech unit candidate group. Consequently, the best speech units are selected rationally and efficiently.

Data supplied from the **esp@cenet** database - I2

⑯ 日本国特許庁 (JP)

⑪ 特許出願公開

⑫ 公開特許公報 (A)

平1-209500

⑩ Int. Cl.

G 10 L 5/04

識別記号

府内整理番号

F-8622-5D

⑬ 公開 平成1年(1989)8月23日

審査請求 未請求 請求項の数 1 (全6頁)

⑤ 発明の名称 音声合成方式

⑭ 特願 昭63-36204

⑮ 出願 昭63(1988)2月17日

特許法第30条第1項適用 昭和62年10月13日社団法人 日本音響学会発行の「日本音響学会 昭和62年度 秋季研究発表会 講演論文集 - I -」に発表

⑯ 発明者 久坂 芳典 京都府相楽郡精華町大字乾谷小字三平谷5番地 株式会社
エイ・ティ・アール自動翻訳電話研究所内

⑰ 出願人 株式会社エイ・ティ・
アール自動翻訳電話研
究所 京都府相楽郡精華町大字乾谷小字三平谷5番地

⑱ 代理人 弁理士 深見 久郎 外2名

明細書

1. 発明の名称

音声合成方式

2. 特許請求の範囲

出力したい音声内容に対応した音韻系列記号信号と韻律制御信号とから任意の語彙の音声合成を行なう方式、および既发声音声の編集により音声合成を行なう方式であって、

予め任意の長さの音韻連接単位からなる音声単位集合を準備し、

前記音韻系列記号信号を受けて、前記音声単位集合から音声合成に使用する可能性のあるすべての音声単位候補群を抽出し、

前記すべての音声単位候補群の中から所定の基準に基づいて使用すべき音声単位を決定する、音声合成方式。

3. 発明の詳細な説明

【産業上の利用分野】

この発明は、音声の合成方式に関するもので、特に、音声単位の編集、結合および変型によって

音声合成を行なう方式に関するものである。

【従来の技術】

現在使われている音声合成方式では、内容の異なった音声を出力するために、音楽、音節、単語および文節といった音声単位を予め蓄えておき、それらの変型や結合により音声を合成する方式を採用している。通常、これらの音声単位は、予め定められた使用箇所に依存して数個用意されるか、あるいは、音韻論的な単位に合わせて1個用意される。しかし、自然な出力音声を得るために、出力したい内容の可変度に応じて、複数個の音声単位が必要であることが指摘されてきた。

そのため、母音・子音・母音 (V C V) 、子音・母音・子音 (C V C) および2音節連鎖 (C V C V) など、長い音声単位を用いることにより隣接する音韻による影響を取り入れた合成方式や、基本周波数 (ピッチ) を多段階に分けて発生した複数の音声単位を用いる合成方式などが考案されてきている。

【発明が解決しようとする課題】

これらの合成方式では、用意されている音声単位の候補数が少ないので、出力したい音声内容に對応した音韻系列記号が与えられた後、音声単位の選択をするために、画一的でかつ局所的な基準が用いられている。したがって、合成された音声の品質は低いものとなっている。品質の高い音声の合成を行なうためには、多くの音声単位数を準備することが必要である。しかし、單に音声単位を増やして種々の音韻構造を持つ多種類の音声単位を音声単位集合として用意しても、従来のこれらの合成方式を用いたのでは、多くの音声単位の中から合理的かつ効率良く適切な音声単位を選出することができないという課題がある。

この発明は、上記のような課題を解決するためになされたもので、多くの音声単位を含む音声単位集合の中から、与えられた選択基準を満たす評価関数に基づき、合理的でかつ効率良く音声単位系列を選択することのできる音声合成方式を提供することを目的とする。

[課題を解決するための手段]

アクセントおよび息継ぎ位置などの韻律制御信号が入力される。入力された音韻系列記号信号は入力音韻系列処理部2に与えられ、入力音韻系列処理部2は入力された音韻系列記号を適当に分割した部分音韻系列を生成する。合成単位エントリファイル4は、音声の合成に必要な種々の音声合成単位情報を予め蓄えておくためのものである。合成単位ラティス生成部3は、入力音韻系列処理部2から出力される部分音韻系列の中から、合成単位として予め蓄えられて存在するものを、合成単位エントリファイル4を検索することによって見い出し、合成に用いられる可能性のある音声合成単位候補のすべてをラティスの形式で列挙して音声テンプレート決定部5に与える。

音声テンプレートファイル7は、音声として出力するのに必要な情報が蓄えられている。音声テンプレート決定部5は、合成単位ラティス生成部3で得られた合成単位ラティスを用い、使用可能な音声テンプレートの組合せのうちから、予め定められた選択基準に基づき最適なものを決定し、

この発明に係る音声合成方式は、予め任意の長さの音韻連接単位からなる音声単位集合を準備しておき、入力された音韻系列記号信号について音声単位集合の中から使用する可能性のあるすべての音声単位候補群を選出し、選出されたすべての音声単位候補群の中から所定の基準に基づいて使用すべき音声単位を決定するものである。

[作用]

この発明における音声合成方式は、音声単位集合の中から入力された音韻系列記号信号について使用する可能性のあるすべての音声単位候補群を予め整理して蓄えておくので、そこから効率良く、使用すべき音声単位を決定することができる。

[発明の実施例]

以下、この発明の実施例を図面を用いて詳細に説明する。

第1図は、この発明による音声合成方式を示すプロック図である。

第1図を参照して、入力端子1には、出力したい音声内容に對応した音韻系列記号信号ならびに

決定された音声テンプレートについて音声テンプレートファイル7から音声として出力するのに必要な情報を得る。この決定方法は、選択基準を選ぶことによって自由に制御することが可能であり、たとえば、合成に要求される音声テンプレートへの条件、すなわち、合成単位ラティス生成部3で得られた合成単位ラティスが持つ出力音韻環境や、韻律制御信号を受けて韻律制御モジュール6が出力する出力韻律パラメータ値と音声テンプレートファイル7に存在する合成用テンプレートの持つ条件（抽出音韻環境および韻律パラメータ値）との整合性や、さらに、音声テンプレート間の接続容易性などをもとに決定する方法である。

テンプレート結合部8は、音声テンプレート決定部5によって決定されたテンプレートを、韻律制御モジュール6によって与えられる値によって伸縮し、テンプレート間の変形および結合処理を行ない、音声合成パラメータ系列を音声合成器9に出力する。音声合成器9は、テンプレート結合部8で得られた音声合成パラメータ系列によって

駆動され、出力端子10に合成された音声波形を出力する。

第2図は、この発明の実施例に用いられる合成単位エントリファイルの一例を示す図である。

この合成単位エントリファイルは、国語辞典内の重要語5240単語の音声を音声テンプレートとして持つ場合に生成した合成単位エントリファイルの一部を示している。このファイルは音声の合成単位を検索する効率を高めるために、各合成単位が構成音韻数ごとに階層化され、木構造を持つ辞書形式に整理されている。第2図は第1ないし第4階層の一部を示したものであり、各合成単位エントリごとに、音韻系列名A、その合成単位テンプレート総数B、テンプレートを含む単語の格納位置（始点C1、終点C2）、一階層下の（すなわち右に1音韻長い）音声単位エントリの総数D、およびそれらの格納位置（始点E1、終点E2）を示す値が蓄えられている。これから、たとえば、合成単位エントリ/a/は全部で3523テンプレート存在し、それらを含む単語はテ

ンプレートファイル中の単語リストの35124番地から75691番地に格納されており、また、/a/を先行音韻として持つ2音韻連接エントリ/aN/、/aa/、…/az/は全部で37個存在し、それらは本エントリファイル中の756番地から1936番地に格納されていることがわかる。

第3図は、この発明の実施例における合成ラティスの作成の様子を示す図である。

第3図では、第2図で示された合成単位エントリファイルを用いて、入力音韻系列記号の「花が咲き出した」/hanagasaki dashita/を入力した場合に作成された合成単位ラティスの一部が示されている。この例では、最初にエントリファイルの検索により/h/から始まる単位エントリとして最長の/hana/が見い出され、その検索過程でより短い/h/、/ha/および/han/の単位エントリが見い出され、第3図に示したようなエントリファイル内の情報を読み出すことができる。この結果、たとえば/h

/ana/は10個のテンプレートを持ち、それは、/hana/（「花」）、/hana/（「鼻」）、…/hanayaka/（「華やか」）などの単語音声サンプルに含まれていることがわかる。

第4図は、この発明の実施例におけるある選択基準を仮定した場合の合成単位ラティスから選択された音声単位とそのテンプレートが含まれる単語音声を示す図である。

第4図では、第3図で示した入力例について、（1）子音から母音への遷移箇所の連続性優先、（2）有聲音連続箇所の連続性優先、（3）長単位優先、（4）単位間の連続性重視などの音声テンプレートを決定するための4つの基準が仮定されている。図に示されるように、最初の/hana/は単語音声サンプル/hanabanashii/（「華々しい」）の語頭から得られ、/g/、/sa/、/ki da/および/shit a/は、それぞれ、/nagasa/（「長さ」）、/murasaki/（「紫」）、/hiki/

/dashii/（「引出し」）および/as hit a/（「明日」）から得られる。これらの基準による選択の具体的実現方法としては、たとえば、Yoshinori Sagisaka、Hiroyasu Sato等により1986年4月に発表された“Word Identification Method for Japanese Text-to-Speech Conversion System”と題された論文（Proc. of ICASSP 86 2411頁-2414頁）に示されるような、各選択基準の優先度に応じた選択のコスト付けを行なうことにより、従来、言語処理における形態素同定法に見られる選択手法等が適用できる。

また、この発明による音声合成方式における音声合成単位の選択方法は、上記で説明したような4つの基準に限られるものではなく、選択基準間の優先度が明確に規定できるものであれば、基準内容、基準個数などによらず適用可能である。

以上で得られた音声テンプレートの影響パラメ

ータ系列は、さらに韻律制御モジュール6から与えられる音韻時間長、振幅および基本周波数などの出力音韻パラメータ値に基づいて、変形が加えられ、結合され、音声合成器8に送られ、合成された音声波形が出力端子10を介して出力される。

以上の説明から明らかなように、この発明の一実施例による音声合成方式は、合成用単位の抽出環境情報を格納した合成単位エントリファイルの使用により、多数個の音声単位列の中から効率良く所望の音声単位列を選択できる利点を有するとともに、選択基準の変更に対しても柔軟に対処できるため、使用する現在の音声合成技術およびそれらの進展に対し、容易に適用できる利点を有する。

また、複数個の任意の長さの音韻接続単位からなる音声単位集合を用いた音声合成を可能とするため、音声単位の結合部で生ずる接続歪および音声単位の抽出環境と使用環境の相違に起因する歪による音声品質の劣化に対して、柔軟に対処することが可能なため、良好な音声品質を持つ音声の

合成が期待できる。

【発明の効果】

以上のように、この発明によれば、予め任意の長さの音韻接続単位からなる音声単位集合を準備し、入力された音韻系列記号信号について音声単位集合の中から使用する可能性のあるすべての音声単位候補群を予め整理して蓄え、その音声単位候補群の中から所定の基準に基づき使用すべき音声単位を決定するようにしたので、使用すべき音声単位を決定するのに、予め蓄えられた音声単位候補群を参照することができ、効率良く使用すべき音声単位を決定することができる。

4. 図面の簡単な説明

第1図は、この発明による音声合成方式を示すブロック図である。第2図は、この発明の実施例に用いられる合成単位エントリファイルの一例を示す図である。第3図は、この発明の実施例における合成ラティスの作成の様子を示す図である。第4図は、この発明の実施例における合成単位ラティスから選択された音声単位とそのテンプレート

トが含まれる単語音声を示す図である。

図において、1は入力端子、2は入力音韻系列処理部、3は合成単位ラティス生成部、4は合成単位エントリファイル、5は音声テンプレート決定部、6は韻律制御モジュール、7は音声テンプレートファイル、8はテンプレート結合部、9は音声合成器、10は出力端子、Aは音韻系列名、Bは合成単位テンプレート総数、C1およびC2はテンプレートを含む単語の格納位置の始点および終点、Dは一階層下の音声単位エントリの総数、E1およびE2は一階層下の音声単位エントリの格納位置の始点および終点である。

特許出願人 株式会社エイ・ティ・アール

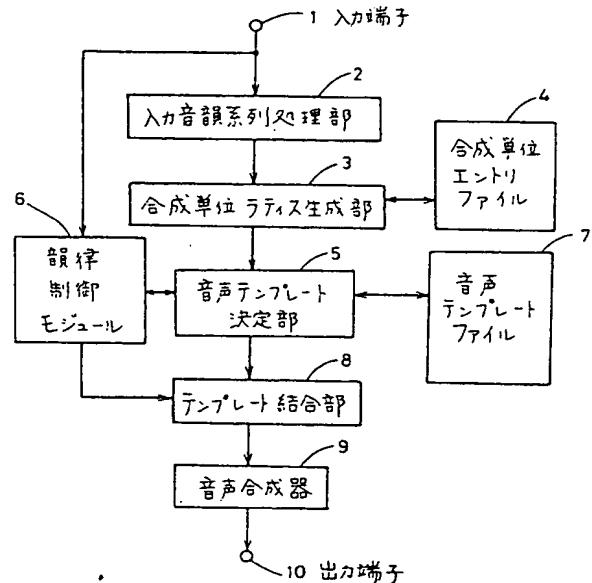
自動翻訳電話研究所

代理人 弁理士 深見久郎

(ほか2名)



第1図



四二

(第 1 層) (第 2 層) (第 3 層) (第 4 層)

A	B	C1	C2	D	E1	E2	F1	F2								
N	11.0	0	1234.9	26	0	155	Ht	J9	J10	3284	J	210	2193			
b	a	352.3	512.0	756.0	317	576.0	I	1330	N	1.5	3352.0	J3707	3	2246	2205	
b	a	152.3	201.0	215.93	5	193.7	I	2102	Nz	2.3	648.9	J3655	3	2306	2425	
bj	a	8.248	30.9	236.0	0.8	1.203	I	213.0	2134	an	2.3	756.92	J7262	23	2126	3122
ch	281	312358	24	0729	4	2135.	2275		za	25	917.04	J5610	3	3130	3104	
cch	9	2367.0	2367.0	2367.0	4	2216	2107		ab	76	877.14	J8855	5	3304	3457	
d	382	25.0	53.5	260.67	7	2.08	251		ach	0	920.70	J316	2	3458	3560	
du	-	1288.87	2688.7	1	2515	2510		acch	1	920.01	J3049	1	3521	3546		
e	1661	2869.17	29788.8	32	2524	1	3633	ad	33	916.6	J9552	3	3575	3630		
e	152.	3513.0	3523.0	3	3634	3355		ae	73	559.02	J5664	6	3633	3821		
e	1524	376.95	391.02	5	3716	3312		af	480.64	J8517	1	3822	3851			
ff	28	29.0002	32.0146	3	3513	4015		ag	18	96.03	J0001	5	3852	4017		
hf	187	3936.01	3936.01	1	4016	1572		ary	103	365.	J0582	1	4018	4084		
hy	19	1.1111	1.1111	1.2	1.1111	2.1158	2226	ch	13	1052.2	J1055	4	4051	4176		
ja	3408	1154.712	471570	38	1227	5317		ei	487	106.07	J1171	1	26173	2656		
j	395	517.0202	5155.515	4	5538	6575		ej	23	121.01	J2144	1	26050	2656		
k	216	513.9312	6222.15	5	5676	5855		ak	43	122.28	J128	1	5069	5162		
ky	81	638.003	689.210	5	5856	6023		akw	28	12.45	J13855	5	5267	5142		
ky	100	7.1012	74.2323	3	6300	6133		aky	2	4788.	J13886	2	5422	5416		
kky	9	639.000	639.015	2	6124	6203		am	12	142.58	J14258	1	5492	5556		
ky	166	756.085	761.075	5	6200	6383		am	12	138.95	J15901	5	5528	5623		
ky	9	8.0120	2.8121	9	5200	6151		an	84	159.22	J15230	5	5701	5874		
n	548	8125.57	8188.01	5	6142	6628		any	10	150.17	J15671	5	5819	5913		
ny	13	8.141.85	8.848.51	2	6229	6597		ao	20	157.58	J15983	7	5914	6014		
p	2160	8.318.85	88.009.05	36	188.05	7135		ap	1	157.52	J17541	1	6145	6178		
pp	26	388.000	38.512.5	7	6175	8.013		app	8	157.02	J17122	2	6179	6282		
py	PY	3.871.6	3675.61	1	8271	8271		appy	1	158.04	J15804	5	6184	6281		
ppj	PY	1.380.78	38.66.63	1	8255	8303		ar	18	158.07	J16258	3	6220	6146		
ry	r	1.61	37.010	5	810.0	8493		as	19	171.95	J17415	1	6463	6100		
ry	81	10.5111	10.501116	1	6500	6803		ass	38	182.67	J18025	4	6504	6744		
s	1083	105824	106619	4	8610	8701		ash	38	170.03	J18120	3	6741	6744		
ss	55	12.511.71	11.162175	1	876.07	8906		ashb	7	182.57	J18145	6	6743	6744		
ss	787	10.906.03	11.08274	1	8207	9061		at	13	186.72	J18704	3	6987	7054		
ss	25	11.262.81	11.631.29	3	9362	9210		atl	12	188.75	J19112	2	7033	7119		
ss	1	865	11.318.88	11.888.23	2	9211	9321		als	0	19.056	J19539	1	7196	72	
ss	55	10.516.55C	12.257.19	5	1251	8.035		alls	2	19.0510	J20443	1	7230	7277		
ss	15	485	12.380.6	12.424.5	1	913.7	9171		au	66	199.25	J20015G	0	0	0	
ss	5	12.670.1	12.271.21	1	915.5	9511		av	67	200.37	J20123	7	2121	2107		
ss	5	17.8	12.731.0	13.220.6	15	3512	1031		av	41	205.01	J20553	3	2108	2107	
ss	1	17.8	12.871.0	13.271.0	15	3512	1031		az	55	207.51	J20712	4	2111	754.8	
ss	1	17.8	12.871.0	13.271.0	15	3512	1031		ba	120	21.5315	J21741	1	754.9	80	
ss	2	22.1154	22.616.9	8	1085	8301		be	42	23.2314	J232314	1	2239.9	80892		

23 DECEMBER

三國志

卷之三

- 739 -

第3図

入力音韻系列記入/hanagasakidashita/(「花が咲き出した」)

合成単位 エントリ名	各エントリの テンプレート数	各エントリを含む 単語音声サンプル
h	487	
ha	207	
han	13	
hana	10	
a	3523	
an	84	
ana	44	
anag	1	
n	548	
na	209	
nag	20	
naga	10	
nagas	3	
nagasa	1	
a	3523	
ag	118	
aga	48	
agas	5	
agasa	1	
g	554	
ga	208	
gas	12	
gasa	1	
a	3523	
as	197	
asa	41	
asak	3	
asaki	1	

第4図

入力音韻系列 hanagasakidashita
↑
① ↑ ② ↑ ③ ↑ ④ ↑ ⑤ ↑

選択された音声単位とそのテンプレートを含む単語音声

- ① / hana banashii / (「華々しい」)
- ② / nagasa / (「長さ」)
- ③ / murasaki / (「紫」)
- ④ / hikidashi / (「引き出レ」)
- ⑤ / ashita / (「明日」)